

1600

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/595,947A

DATE: 11/08/2001
TIME: 07:45:44

Input Set : A:\ST96042Asqlt.txt
Output Set: N:\CRF3\11082001\I595947A.raw

3 <110> APPLICANT: ICARD-LIEPKALNS, Christine
4 MALLET, Jacques
5 RAVASSARD, Philippe
7 <120> TITLE OF INVENTION: POLYPEPTIDES OF THE "BASIC-HELIX-LOOP-HELIX" bHLH
8 FAMILY, CORRESPONDING NUCLEIC ACID SEQUENCES
10 <130> FILE REFERENCE: ST96042A-US
12 <140> CURRENT APPLICATION NUMBER: US 09/595,947A
13 <141> CURRENT FILING DATE: 2000-06-16
15 <150> PRIOR APPLICATION NUMBER: FR96/15651 *OK*
16 <151> PRIOR FILING DATE: 1996-12-19 *b/c*
18 <150> PRIOR APPLICATION NUMBER: PCT/FR97/02368 *Erroneous: See next*
19 <151> PRIOR FILING DATE: 1997-12-19
21 <160> NUMBER OF SEQ ID NOS: 28
23 <170> SOFTWARE: PatentIn Ver. 2.1
25 <210> SEQ ID NO: 1
26 <211> LENGTH: 1460
27 <212> TYPE: DNA
28 <213> ORGANISM: Rattus norvegicus
30 <400> SEQUENCE: 1
31 gcaggttagcg agaggagcag tccctggcc cccgttgctg attggcccg ggcacaggca 60
32 gcagcccgcc aggcacgctc ctggtcggg cagagcagat aaagcgtgcc aggggacaca 120
33 cgattagcag ctcagaagtc cctctgggtc tcaccactgc acagaggccg aggacccct 180
34 ccgagcttct ttgctgcctc cagacgcaat ttactccagg cgagggcgcc tgcagcttag 240
35 caaaaacttcg aagcgacgag aggggttcag ctatccaccc ctgttctact ctgaccaccc 300
36 gcagctctct gtttttga gcccggagta actagtaac atttaggaac ctccaaaggg 360
37 tagaagaggg gagtgggtgg gcttactcta gtcccgctg gatgtaccc taagtcagag 420
38 actgtcacac ccccttcca ttttttccca acctcaggat ggccctcat cccttggatg 480
39 cccccaccat ccaagtgtcc caagagaccc agcaaccctt tcccgagcc tcggaccacg 540
40 aagtgttcag ttccaaattcc accccaccta gcccactct cgtaccgagg gactgtccg 600
41 aagcagaagc aggtgactgc cgagggacat cgaggaagct ccgtgcgcgg cgccggaggc 660
42 gcaacaggcc caagagcggag ttggcactga gcaaggcgg acgaagccgg cgcaagaagg 720
43 ccaacgaccg ggagcgaac cgcattgcaca accttaactc cgcgttggat ggcgtgcgc 780
44 gtgtcctgcc cacccccc gatgacgcca aacttacaaa gatcgagacc ctgcgttccg 840
45 cccacaacta catttggca ctgactcaga cgctgcgcatt agccggaccac agcttctacg 900
46 gccccgagcc ccctgtgccc tggggggagc tgggaagccc gggagggggc tccagcggcg 960
47 actggggctc tatctactcc ccagttccc aagctggtag cctgagccccc acagcctcat 1020
48 tggaggagtt ccctggctg caggtgcccc gctcccccattt ctgtctgtcc ccgggcaccc 1080
49 tgggtttctc agacttcttg tgaagggccc aaacaggccc tgggcgggtgg ggcgtggcag 1140
50 aaaggggaggg agtcagagct gtctgaaatg gaaggtagtg gaggcactcg agcatctcgc 1200
51 cccttctggc tttcattagt caggtccctg atttaaccag gattgcacca gttccttgct 1260
52 gctgtgcgtg cacaaggac attgcaggct gatctcttct taaccctct cagtgcc 1320
53 acctcaaact cccgtccaa gcagaggaga gccgtacac taaatagtt ggagactccc 1380
54 atacttcctg gtgactccgc cctcttcaa atctgcgggc ctccaaaccac cgcttctcc 1440
55 agagtgacct aatccagtgt 1460
58 <210> SEQ ID NO: 2
59 <211> LENGTH: 24
60 <212> TYPE: DNA

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/595,947A

DATE: 11/08/2001
TIME: 07:45:44

Input Set : A:\ST96042Asqlt.txt
Output Set: N:\CRF3\11082001\I595947A.raw

61 <213> ORGANISM: Artificial Sequence
63 <220> FEATURE:
64 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
66 <220> FEATURE:
67 <223> OTHER INFORMATION: n = Inosine *✓*
69 <400> SEQUENCE: 2
W--> 70 aatk~~h~~gm^gagcgcndkcg cryg
73 <210> SEQ ID NO: 3
74 <211> LENGTH: 24
75 <212> TYPE: DNA
76 <213> ORGANISM: Artificial Sequence
78 <220> FEATURE:
79 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primers
81 <400> SEQUENCE: 3
82 ggcsrdtytc aggtsybga yctt
85 <210> SEQ ID NO: 4
86 <211> LENGTH: 25
87 <212> TYPE: DNA
88 <213> ORGANISM: Artificial Sequence
90 <220> FEATURE:
91 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primers
93 <400> SEQUENCE: 4
94 aaccttaact ccgcgtcgg tgcgc
97 <210> SEQ ID NO: 5
98 <211> LENGTH: 18
99 <212> TYPE: DNA
100 <213> ORGANISM: Artificial Sequence
102 <220> FEATURE:
103 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primers
105 <400> SEQUENCE: 5
106 cgcgggttcc tgccacc
109 <210> SEQ ID NO: 6
110 <211> LENGTH: 6
111 <212> TYPE: DNA
112 <213> ORGANISM: Artificial Sequence
114 <220> FEATURE:
115 <223> OTHER INFORMATION: Description of Artificial Sequence: E box
117 <400> SEQUENCE: 6
118 caggtg
121 <210> SEQ ID NO: 7
122 <211> LENGTH: 6
123 <212> TYPE: DNA
124 <213> ORGANISM: Artificial Sequence
126 <220> FEATURE:
127 <223> OTHER INFORMATION: Description of Artificial Sequence: Mutated E box
129 <400> SEQUENCE: 7
130 tccgtg
133 <210> SEQ ID NO: 8
134 <211> LENGTH: 214

FYI: It is preferred that you describe your unknowns in fields 24, 221 as well as 222 and 223.

24

25

18

6

6

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/595,947A

DATE: 11/08/2001
TIME: 07:45:44

Input Set : A:\ST96042Asqlt.txt
Output Set: N:\CRF3\11082001\I595947A.raw

135 <212> TYPE: PRT
 136 <213> ORGANISM: Rattus norvegicus
 138 <400> SEQUENCE: 8
 139 Met Ala Pro His Pro Leu Asp Ala Pro Thr Ile Gln Val Ser Gln Glu
 140 1 5 10 15
 142 Thr Gln Gln Pro Phe Pro Gly Ala Ser Asp His Glu Val Leu Ser Ser
 143 20 25 30
 145 Asn Ser Thr Pro Pro Ser Pro Thr Leu Val Pro Arg Asp Cys Ser Glu
 146 35 40 45
 148 Ala Glu Ala Gly Asp Cys Arg Gly Thr Ser Arg Lys Leu Arg Ala Arg
 149 50 55 60
 151 Arg Gly Gly Arg Asn Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
 152 65 70 75 80
 154 Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
 155 85 90 95
 157 His Asn Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
 158 100 105 110
 160 Phe Pro Asp Asp Ala Lys Leu Thr Lys Ile Glu Thr Leu Arg Phe Ala
 161 115 120 125
 163 His Asn Tyr Ile Trp Ala Leu Thr Gln Thr Leu Arg Ile Ala Asp His
 164 130 135 140
 166 Ser Phe Tyr Gly Pro Glu Pro Pro Val Pro Cys Gly Glu Leu Gly Ser
 167 145 150 155 160
 169 Pro Gly Gly Gly Ser Ser Gly Asp Trp Gly Ser Ile Tyr Ser Pro Val
 170 165 170 175
 172 Ser Gln Ala Gly Ser Leu Ser Pro Thr Ala Ser Leu Glu Glu Phe Pro
 173 180 185 190
 175 Gly Leu Gln Val Pro Ser Ser Pro Ser Cys Leu Leu Pro Gly Thr Leu
 176 195 200 205
 178 Val Phe Ser Asp Phe Leu
 179 210
 182 <210> SEQ ID NO: 9
 183 <211> LENGTH: 1330
 184 <212> TYPE: DNA
 185 <213> ORGANISM: Homo sapiens
 187 <400> SEQUENCE: 9
 188 cctcggacc cattctct tctttctcc tttggggctg gggcaactcc caggcgaaaa 60
 189 cgcctgcgc tcagctgaac ttggcgacca gaagcccgct gagctccca cggccctcgc 120
 190 tgctcatgc tctctattct ttgcgcgg tagaaaggta atatttggag gccttcgagg 180
 191 gacgggcagg ggaaagaggg atcctctgac ccagcgaaaa ctgggaggat ggctgtttt 240
 192 gtttttccc acctagcctc gaaatcgccg actgcgcgt gacggactca aacttaccct 300
 193 tccctctgac cccgcgttag gatgacgcct caaccctcgg gtgcgcacac tgcggccatg 360
 194 acccgtgaga cggagcggtc cttccccaga gcctcgaaag acgaagtgac ctgcacccacg 420
 195 tccgccccgc ccagccccac tcgcacacccg gggaaactgcg cagaggcgga agagggaggc 480
 196 tgccgagggg ccccgaggaa gctccggca cggcgaaaa gacgcagccg gcctaagac 540
 197 gagttggcac tgagcaagca gcgacggagt cggcgaaaaga aggccaacga ccgcgagcgc 600
 198 aatcgaatgc acgaccta ctcggactg gacgcctgc gcggtgtcct gcccacctc 660
 199 ccagacgacg cgaagctcac caagatcgag acgctgcgt tcgcccacaa ctacatctgg 720
 200 gcgctgactc aaacgctgcg catagcggac cacagcttgt acgctgcgt gcccggcg 780

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/595,947A

DATE: 11/08/2001
TIME: 07:45:44

Input Set : A:\ST96042Asqlt.txt
Output Set: N:\CRF3\11082001\I595947A.raw

201 ccgcactgcg gggagctggg cagcccaggc ggtccccccg gggactgggg gtcctctac 840
 202 tccccagtct cccaggctgg cagcctgagt cccggcggt cgctggagga gcgacccggg 900
 203 ctgctgggg ccacctcttc cgcctgctt agccaggca gtctggctt ctcagatttt 960
 204 ctgtgaaagg acctgtctgt cgctgggtg tgggtgctaa gggtaaggga gagggaggg 1020
 205 gccgggagcc gtagagggtg gccgacggcg gccccctca aaagcactt ttccttctgc 1080
 206 ttctccctag ctgaccctg gccccccag gcctccacgg gggcggtttagg ctgggttcat 1140
 207 tccccggccc tccgagccgc gccaacgcac gcaaccctt ctgtgccc cgcaagtgg 1200
 208 gcattgcaaa gtgcgtcat tttaggcctc ctctctgcca ccacccata atccattca 1260
 209 aagaatacta gaatggtagc actacccggc cggagccgccc caccgttgg ggtcgcccta 1320
 210 ccctcactca 1330
 213 <210> SEQ ID NO: 10
 214 <211> LENGTH: 214
 215 <212> TYPE: PRT
 216 <213> ORGANISM: Homo sapiens
 218 <400> SEQUENCE: 10
 219 Met Thr Pro Gln Pro Ser Gly Ala Pro Thr Val Gln Val Thr Arg Glu
 220 1 5 10 15
 222 Thr Glu Arg Ser Phe Pro Arg Ala Ser Glu Asp Glu Val Thr Cys Pro
 223 20 25 30
 225 Thr Ser Ala Pro Pro Ser Pro Thr Arg Thr Pro Gly Asn Cys Ala Glu
 226 35 40 45
 228 Ala Glu Glu Gly Gly Cys Arg Gly Ala Pro Arg Lys Leu Arg Ala Arg
 229 50 55 60
 231 Arg Gly Gly Arg Ser Arg Pro Lys Ser Glu Leu Ala Leu Ser Lys Gln
 232 65 70 75 80
 234 Arg Arg Ser Arg Arg Lys Lys Ala Asn Asp Arg Glu Arg Asn Arg Met
 235 85 90 95
 237 His Asp Leu Asn Ser Ala Leu Asp Ala Leu Arg Gly Val Leu Pro Thr
 238 100 105 110
 240 Phe Pro Asp Asp Ala Lys Leu Thr Lys Ile Glu Thr Leu Arg Phe Ala
 241 115 120 125
 243 His Asn Tyr Ile Trp Ala Leu Thr Gln Thr Leu Arg Ile Ala Asp His
 244 130 135 140
 246 Ser Leu Tyr Ala Leu Glu Pro Pro Ala Pro His Cys Gly Glu Leu Gly
 247 145 150 155 160
 249 Ser Pro Gly Gly Pro Pro Gly Asp Trp Gly Ser Leu Tyr Ser Pro Val
 250 165 170 175
 252 Ser Gln Ala Gly Ser Leu Ser Pro Ala Ala Ser Leu Glu Glu Arg Pro
 253 180 185 190
 255 Gly Leu Leu Gly Ala Thr Ser Ser Ala Cys Leu Ser Pro Gly Ser Leu
 256 195 200 205
 258 Ala Phe Ser Asp Phe Leu
 259 210
 262 <210> SEQ ID NO: 11
 263 <211> LENGTH: 18
 264 <212> TYPE: DNA
 265 <213> ORGANISM: Artificial Sequence *DR*
 267 <220> FEATURE:
 268 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer *PF*

RAW SEQUENCE LISTING
PATENT APPLICATION: US/09/595,947A

DATE: 11/08/2001
TIME: 07:45:44

Input Set : A:\ST96042Asqlt.txt
Output Set: N:\CRF3\11082001\I595947A.raw

270 <400> SEQUENCE: 11
271 caacgaccgg cagcgcaa
274 <210> SEQ ID NO: 12
275 <211> LENGTH: 24
276 <212> TYPE: DNA
277 <213> ORGANISM: Artificial Sequence
279 <220> FEATURE:
280 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
282 <400> SEQUENCE: 12
283 gcccgatgt agttgtggc gaag
286 <210> SEQ ID NO: 13
287 <211> LENGTH: 60
288 <212> TYPE: DNA
289 <213> ORGANISM: Artificial Sequence
291 <220> FEATURE:
292 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
294 <400> SEQUENCE: 13
W--> 295 atcgttgaga ctcgtaccag cagagtacg agagagacta cacggtaactg gnnnnnnnn 60
298 <210> SEQ ID NO: 14
299 <211> LENGTH: 20
300 <212> TYPE: DNA
301 <213> ORGANISM: Artificial Sequence
303 <220> FEATURE:
304 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
306 <400> SEQUENCE: 14
307 agacgacgca gaaatccatca
310 <210> SEQ ID NO: 15
311 <211> LENGTH: 24
312 <212> TYPE: DNA
313 <213> ORGANISM: Artificial Sequence
315 <220> FEATURE:
316 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
318 <400> SEQUENCE: 15
319 gctcaccaag atcgagacgc tgcg
322 <210> SEQ ID NO: 16
323 <211> LENGTH: 25
324 <212> TYPE: DNA
325 <213> ORGANISM: Artificial Sequence
327 <220> FEATURE:
328 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
330 <400> SEQUENCE: 16
331 atcgttgaga ctcgtaccag cagag
334 <210> SEQ ID NO: 17
335 <211> LENGTH: 25
336 <212> TYPE: DNA
337 <213> ORGANISM: Artificial Sequence
339 <220> FEATURE:
340 <223> OTHER INFORMATION: Description of Artificial Sequence: PCR Primer
342 <400> SEQUENCE: 17

24
No possible values are given
for n. No location information
is given for n. Therefore Erroneous

Must include fields
221, 222 and 223 responses
to explain nucleotide presence.

good

20

24

25

VERIFICATION SUMMARY

PATENT APPLICATION: US/09/595,947A

DATE: 11/08/2001

TIME: 07:45:45

Input Set : A:\ST96042Asqlt.txt

Output Set: N:\CRF3\11082001\I595947A.raw

L:70 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:2
L:70 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:2
L:70 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:2
L:295 M:258 W: Mandatory Feature missing, <221> not found for SEQ ID#:13
L:295 M:258 W: Mandatory Feature missing, <222> not found for SEQ ID#:13
L:295 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:13